Distributions of Genotypes after a Panmixture

By Yusaku Komatu

(Received Feb. 19, 1954)

1. Introduction.

Consider a population of size $2N$ consisting of $N$ females and $N$ males. We observe a single inherited character which consists of $m$ multiple alleles at one diploid locus denoted by

$$A_i \quad (i=1, \ldots, m)$$

and of which the inheritance is subject to Mendelian law.

There are $m(m+1)/2$ possible genotypes $A_A A_B (a,b=1,\ldots,m; a \leq b)$ among which $m$ types $A_A A_B (b=1,\ldots,m)$ are homozygous and $m(m-1)/2$ types $A_A A_B (a,b=1,\ldots,m; a < b)$ are heterozygous. Let the distributions of these $m(m+1)/2$ genotypes $A_A A_B (a \leq b)$ in females and in males be designated by

$$F_{ab} \quad \text{and} \quad M_{ab} \quad (a,b=1,\ldots,m; a \leq b)$$

or, as the aggregates, by

$$\mathcal{F} = (F_{11}, \ldots, F_{mm}, F_{12}, \ldots, F_{m-1,m})$$

and

$$\mathcal{M} = (M_{11}, \ldots, M_{mm}, M_{12}, \ldots, M_{m-1,m})$$

respectively, so that

$$\sum_{a<b} F_{ab} = \sum_{a<b} M_{ab} = N.$$ 

The order of genes in a genotype being immaterial, both genotypes $A_A A_B$ and $A_B A_A$ are regarded as identical each other even when the suffices $a$ and $b$ are distinct. Accordingly, we put $F_{ab} = F_{ba}$ and $M_{ab} = M_{ba}$.

We now introduce a set of stochastic variables

$$\mathcal{C} = (C_{11}, \ldots, C_{mm}, C_{12}, \ldots, C_{m-1,m})$$
Distributions of Genotypes after a Panmixia

extending over non-negative integers and satisfying a single equation

$$\sum_{a\leq b} C_{ab} = N.$$ 

We designate by

$$\mathcal{V}(\mathcal{C}) = \mathcal{V}(\mathcal{C} | \mathcal{Z}; \mathcal{M})$$

the probability that a distribution in the next generation after a panmixia is given by $\mathcal{C}$, each mating being supposed to produce one child so that, as stated above, $\sum_{a\leq b} C_{ab} = N$.

In order to preserve the size of the whole population in the next generation, it would be rather preferable to suppose that each mating produces two children, one female and one male. In the present purpose, however, one may think to confine oneself to a distribution in either sex in the next generation, by supposing that the fertility of every mating is equal.

The main purpose of the present paper is to establish an explicit expression for a probability-generating function $\Phi$ defined by

$$\Phi(\mathcal{\hat{Z}}) = \Phi(\mathcal{\hat{Z}} | \mathcal{Z}; \mathcal{M}) = \sum_{\mathcal{C}} \mathcal{V}(\mathcal{C} | \mathcal{Z}; \mathcal{M}) \prod_{a\leq b} C_{ab},$$

$\mathcal{\hat{Z}}$ designating a set of indeterminate variables:

$$\mathcal{\hat{Z}} = (z_{11}, \ldots, z_{mm}, z_{12}, \ldots, z_{m-1,m});$$

the summation extends over the whole range of $\mathcal{C}$.

By the way, it is noted that a variable involved in $\mathcal{C}$, for instance $C_{m-1,m}$, similarly a variable involved in $\mathcal{Z}$ and in $\mathcal{M}$, and consequently a variable involved in $\mathcal{\hat{Z}}$, for instance $z_{m-1,m}$, may be omitted, since we have supposed that total numbers of individuals are fixed in children, in females, and in males. The omission of $z_{m-1,m}$ corresponds, of course, to put $z_{m-1,m} = 1$. The range of summation in the last expression is then replaced by the sets of $m(m+1)/2 - 1$ non-negative integers $C_{ab}$ ($a, b = 1, \ldots, m; a \leq b; (a, b) \neq (m-1, m)$) satisfying

$$\sum_{a\leq b} C_{ab} - C_{m-1,m} \leq N.$$ 

However, we shall retain, for the sake of formal symmetry, the dependent quantities concerning a genotype $A_{m-1}A_m$ in our expressions, since they will give rise to few complications.
We now state a further supplementary remark. Though our original problem concerns \( m \) different genes, it is reducible to a problem with a lower \( m \), according to circumstances, for instance, when there concern routine statistics such as means, variances, covariances, etc. on the stochastic variables under consideration. In fact, if we consider a homozygous type \( A_iA_i \), the distinction among the genes except \( A_i \) becomes a matter of indifference. Consequently, if we put

\[
\sum_{b \neq i} F_{ib} = F_{iw}, \quad \sum_{a, b \neq i} F_{ab} = F_{ww},
\]
\[
\sum_{b \neq i} M_{ib} = M_{iw}, \quad \sum_{a, b \neq i} M_{ab} = M_{ww},
\]
\[
\sum_{b \neq i} C_{ib} = C_{iw}, \quad \sum_{a, b \neq i} C_{ab} = C_{ww},
\]

then the mean and variance of the random variable \( C_{ii} \) in the distributions in the next generation coincide really with those which result from considering only two genes, i.e. the gene \( A_i \) itself and an imaginary gene \( A_w \) introduced in a manner stated just above. On the other hand, if we consider a heterozygous type \( A_iA_j \), the distinction among the genes except \( A_i \) and \( A_j \) becomes a matter of indifference. Consequently, if we put

\[
\sum_{b \neq i, j} F_{ib} = F_{iw}, \quad \sum_{b \neq i, j} F_{jb} = F_{jw}, \quad \sum_{a, b \neq i, j} F_{ab} = F_{ww}, \quad \text{etc.,}
\]

the mean and variance of the random variable \( C_{ij} \) coincide just with those which result from considering only three genes, i.e. the genes \( A_i \) and \( A_j \) themselves and an imaginary gene \( A_w \). Quite similarly, in order to determine the covariance between any two genotypes, it suffices to study a problem in which there concern at most four genes.

In spite of such being the case, and moreover though the arguments for any \( m \) can be performed, in principle, quite similarly to those for such lower \( m \), for instance \( m=2, 3 \) or 4, we shall treat in the following lines the original problem itself for the sake of completeness.

In a series of previous papers\(^1\), we have dealt with analogous

Distributions of Genotypes after a Panmixia

problems far extensively but with few preciseness. We have considered there a population of infinite size and studied merely the means of several distributions. The results of the present paper will give generalizations of some of the previous results.


We now observe \( m \) distinct genes \( A_i \) \((i=1, \cdots, m)\). We first consider a partition of \( N \) males into \( m(m+1)/2 \) classes according to the same number of genotypes \( A_aA_b \) \((a,b=1, \cdots, m; a \leq b)\) of females to be married. Namely, let each of \( M_{ab} \) \((a,b=1, \cdots, m; a \leq b)\) individuals in male-population be divided into \( m(m+1)/2 \) classes, here empty classes being admissible, in such a manner

\[
M_{ab} = \sum_{c \leq d} x_{abcd}, \quad (a,b,c,d=1, \cdots, m),
\]

\[
F_{cd} = \sum_{a \leq b} x_{abcd}, \quad (a \leq b; c \leq d).
\]

Let the matings take place such that, for every pair of \( c \) and \( d \) with \( c, d=1, \cdots, m; c \leq d \), \( x_{abcd} \) \((a,b=1, \cdots, m; a \leq b)\) males of genotypes \( A_aA_b \) are combined, as a whole, with \( F_{cd} \) females of genotype \( A_cA_d \).

All the possible permutations of \( N \) males consisting of \( M_{ab} \) \((a,b=1, \cdots, m; a \leq b)\) individuals of genotypes \( A_aA_b \), respectively, then amount to

\[
N! / \prod_{a \leq b} M_{ab}!
\]

while the permutations of \( F_{cd} \) males to be married with females of genotype \( A_cA_d \), these males consisting of \( x_{abcd} \) \((a,b=1, \cdots, m; a \leq b)\) individuals of genotypes \( A_aA_b \), respectively, amount to

\[
F_{cd}! / \prod_{a \leq b} x_{abcd}!.
\]

On the other hand, it is supposed that any mating \( A_aA_b \times A_cA_d \) produces each of four possible genotypes \( A_aA_c, A_aA_d, A_bA_c \) and \( A_bA_d \) equally likely, that is, with probability \( 1/4 \). When some of genes are coincident and hence two or four of the genotypes to be produced are identical, the probability is then, of course, interpreted as the corresponding sum, namely, as \( 1/2 \) or \( 1 \), respectively. To state more
precisely, if we denote by $i, j, h$ and $k$ the suffices indicating the genes different each other, then the matings

$$A_iA_i \times A_iA_i, \quad A_iA_i \times A_iA_k, \quad A_iA_i \times A_kA_k, \quad A_iA_i \times A_hA_k,$$

$$A_iA_j \times A_iA_j, \quad A_iA_j \times A_iA_k, \quad A_iA_j \times A_hA_k$$

produce a child of type $A_iA_i; A_jA_j; A_iA_k; A_iA_h; A_jA_k; A_jA_h$ with probabilities

$$1, \quad 1/2, \quad 0, \quad 0, \quad 1/4, \quad 1/4, \quad 0;$$

$$0, \quad 0, \quad 0, \quad 0, \quad 1/4, \quad 0, \quad 0;$$

$$0, \quad 0, \quad 0, \quad 0, \quad 1/2, \quad 1/4, \quad 0;$$

$$0, \quad 1/2, \quad 1, \quad 1/2, \quad 0, \quad 1/4, \quad 1/4;$$

$$0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 1/4;$$

$$0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 1/4, \quad 1/4,$$

respectively.

Consequently, the generating function is given by

$$\Phi(3) = \sum_{j \in S} \Phi(j) \sum_{f \in S} z_{j,f}^D = \sum_{j \in S} \frac{\prod_{a \leq b} M_{ab}!}{N!} \sum_{k \leq d} \frac{\prod_{a \leq b} F_{cd}!}{d!} x_{abcd}$$

$$\cdot \sum_{a \leq c} x_{abc} \prod_{a \leq b \leq c} \frac{x_{abc}}{x_{abc}!} \frac{x_{abc}}{x_{abc}!} - \frac{1}{2} \left( \frac{z_{ac}}{2} \right) \frac{x_{abc}}{x_{abc}!} \frac{x_{abc}}{x_{abc}!}$$

$$\cdot \prod_{a \leq c \leq d} \frac{x_{abcd}!}{x_{abcd}!} \frac{x_{abcd}!}{x_{abcd}!} \left( \frac{z_{ac}}{2} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{ad}}{2} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{ab}}{2} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{bb}}{4} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{bb}}{4} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{bd}}{4} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{bd}}{4} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!} \left( \frac{z_{bd}}{4} \right) \frac{x_{abcd}}{x_{abcd}!} \frac{x_{abcd}}{x_{abcd}!}$$

where we make an agreement

$$z_{fg} = z_{gf}, \quad x_{abcd} = x_{abcd}.$$
Distributions of Genotypes after a Panmixia

In the last expression, we further put

$$\sum_{f=g} x^{(f)}_{abcd} = x_{abcd}$$

in which the summation runs over possible types $A_f A_g$ of a child originated from a mating $A_a A_b \times A_c A_d$. In case where a mating $A_a A_b \times A_c A_d$ can produce only one type of children, namely, when $A_a A_b$ and $A_c A_d$ both are homozygous, i.e. $a=b$ and $c=d$, the quantity $x^{(f)}_{abcd}$ is interpreted to be equal to $x_{abcd}$, i.e. $x^{(a)}_{abcd} = x_{abcd}$. Among the quantities here concerned there exist further several relations expressing their dependency. They may be set out as follows:

$$\sum_{a<b, c<d} x_{abcd}^{(f)} = C_{fg} \quad (f, g = 1, \ldots, m; f \leq g),$$

$$\sum_{c<d} x_{abcd} = M_{ab}, \quad \sum_{a<b} x_{abcd} = F_{cd} \quad (a, b, c, d = 1, \ldots, m; a \leq b; c \leq d).$$

Taking these relations of dependence into account, the last summation in the above expression for $\phi(5)$ extends over all the possible sets of non-negative integers $\{x_{abcd}^{(f)}\}$, the second summation over all the possible partitions represented by the sets $\chi = \{x_{abcd}\}$, and the first summation over all the sets of non-negative integers $\{C_{fg}\}$ with total sum $N$.

We first perform, by making use of multinomial identities, the summation with respect to suffices $fg$ involved in $\{x_{abcd}\}$ and $\{C_{fg}\}$, whence follows a relation

$$\phi(3) = \frac{\prod_{a<b} M_{ab}!}{N!} \sum_{\chi} \prod_{C_{fg}} \frac{F_{cd}!}{x_{abcd}^{(f)}!} \prod_{\alpha < \beta} \left(\frac{\alpha_{ab} + \alpha_{ac} + \alpha_{ad}}{2}\right)^{x_{abc}} \prod_{\beta < \gamma} \left(\frac{\beta_{ab} + \beta_{bc} + \beta_{bd}}{2}\right)^{x_{abcd}}$$

$$\cdot \prod_{\alpha < \beta < \gamma} \left(\frac{\alpha_{ab} + \beta_{ab} + \gamma_{ab}}{4}\right)^{x_{abab}} \prod_{\alpha < \beta < \gamma < \delta} \left(\frac{\alpha_{ab} + \beta_{ab} + \gamma_{ab} + \delta_{ab}}{4}\right)^{x_{abcd}},$$

which may be written in a more brief form

$$\phi(3) = \frac{\prod_{a<b} M_{ab}!}{N!} \sum_{\chi} \prod_{C_{fg}} \frac{F_{cd}!}{x_{abcd}^{(f)}!} \prod_{a < b} \left(\frac{\alpha_{ab} + \beta_{ab} + \gamma_{ab}}{4}\right)^{x_{abcd}}.$$
We now introduce, with parameters 
\[ t = (t_{ab}) \quad (a, b = 1, \ldots, m; a \leq b), \]
a function defined by
\[
\Phi(t|t) = \Phi(t|s; M) = \frac{1}{N} \prod_{a \leq b} \frac{M_{ab}}{t_{ab}^M} \cdot \prod_{c < d} \left( \sum_{a \leq b} t_{ab} \frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^{F_{cd}}.
\]

Then, by remembering the relations
\[
\sum_{a \leq b} x_{abcd} = F_{cd}, \quad \sum_{c < d} x_{abcd} = M_{ab},
\]
the multinomial theorem implies that our generating function \( \Phi(t) \) is given by the constant term, i.e. the coefficient of the term \( \prod_{a \leq b} t_{ab}^0 \) in the Laurent expansion around the origin of the last quantity \( \Phi(t|t) \) regarded as a rational function of \( m(m+1)/2 \) variables \( t = (t_{ab}) (a, b = 1, \ldots, m; a \leq b) \).

We further introduce, for a later purpose, with parameters 
\[ s = (s_{ab}) \quad (a, b = 1, \ldots, m; a \leq b), \]
a function defined by
\[
\Phi(s|t; s) = \Phi(s|t; s; M) = \frac{1}{N^2} \prod_{a \leq b} \frac{F_{ab}}{s_{ab}^F M_{ab}^M} \cdot \left( \sum_{a \leq b, c < d} s_{ab} t_{cd} \frac{z_{ac} + z_{ad} + z_{bc} + z_{bd}}{4} \right)^N.
\]

Then, by taking into account the relation
\[
\sum_{c < d} F_{cd} = N,
\]
the multinomial theorem again implies that our generating function \( \Phi(s) \) is given also by the constant term in the Laurent expansion around the origin of the last quantity \( \Phi(s|t; s) \) regarded as a rational function of \( m(m+1) \) variables \( s = (s_{ab}) \) and \( t = (t_{ab}) (a, b = 1, \ldots, m; a \leq b) \).

By separating the factors involved in the last two functions according to ones concerning homozygous and heterozygous types, they are brought into the following forms:
Distributions of Genotypes after a Panmixia

\[ \Phi(\mathbf{\beta} \mid \mathbf{t}) = \frac{1}{N!} \prod_i M_{ii}! \prod_{i < j} M_{ij}! \cdot \prod_i \left( t_{ii} z_{ii} + \sum_{b \neq i} \left( t_{bb} z_{ib} + t_{ib} \frac{z_{ii} + z_{ib}}{2} \right) + \sum_{a,b \neq i} t_{ab} \frac{z_{ia} + z_{ib}}{2} \right)^{F_{ii}} \cdot \prod_{i < j} \left( t_{ii} \frac{z_{ii} + z_{ij}}{2} + t_{jj} \frac{z_{jj} + z_{ij}}{2} + t_{ij} \frac{z_{ii} + 2z_{ij} + z_{jj}}{4} \right) \cdot \sum_{b \neq i,j} \left( t_{bb} \frac{z_{ib} + z_{jb}}{2} + t_{ib} \frac{z_{ii} + z_{ij} + z_{ib} + z_{jb}}{4} + t_{ib} \frac{z_{jj} + z_{ib} + z_{jb} + z_{ij}}{4} \right)^{F_{ij}} \cdot \sum_{a,b \neq i,j} t_{ab} \frac{z_{ia} + z_{ib} + z_{ja} + z_{jb}}{4}^{F_{ij}} \]

and

\[ \mathbf{\Phi}(\mathbf{\beta} \mid \mathbf{\delta}; t) = \frac{1}{N!^2} \prod_i \frac{F_{ii}! M_{ii}!}{s_{ii}! M_{ii}!} \prod_{i < j} \frac{F_{ij}! M_{ij}!}{s_{ij}! M_{ij}!} \cdot \left( \sum_i \left( u_i v_i z_{ii} + \sum_{k < j} (u_i v_j + u_j v_i) z_{ij} \right) \right)^N, \]

where we put, for the sake of brevity,

\[ u_i = s_{ii} + \sum_{b \neq i} \frac{s_{ib}}{2}, \quad v_i = t_{ii} + \sum_{b \neq i} \frac{t_{ib}}{2} \]  

(i = 1, ..., m).

It would be noticed, in passing, that the generating function \( \Phi(\beta) \) itself is expressible, for instance, by means of a contour integral in a form

\[ \Phi(\mathbf{\beta}) = \frac{1}{(2\pi \sqrt{-1})^{m(m+1)/2}} \int \cdots \int \mathbf{\Phi}(\mathbf{\beta} \mid \mathbf{\delta}; t) \prod_{a \neq b} ds_{ab} dt_{ab}, \]

where the multiple integration is taken along the unit circumferences

\[ |s_{ab}| = 1, \quad |t_{ab}| = 1 \]  

(a, b = 1, ..., m; a ≤ b)

in the positive sense on respective complex planes. However, this integral representation will not be availed in the following lines.


By virtue of its own meaning, there must hold an identity
\[ \phi(\xi) = 1, \]

where we put
\[ \xi = \left( \frac{1}{1}, \ldots, \frac{m(m+1)^{1/2}}{1} \right). \]

It is of course that our expression derived above conforms to this requirement. In fact, since we have
\[ \Phi(\xi) \equiv \text{Fab} \equiv M_{ab} \equiv N \left( \sum_{a,b} s_{ab} \sum_{a,b} t_{ab} \right)^{N-1}, \]
we get, by picking up the constant terms in the Laurent expansion,
\[ \phi(\xi) = \frac{1}{N!} \prod_{a,b} \text{Fab} \equiv M_{ab} \equiv N \left( \sum_{a,b} s_{ab} \sum_{a,b} t_{ab} \right)^{N-1} = 1. \]

Now, an analytical expression for the generating function having been established in an explicit manner, the mean of every stochastic variable \( C_{ab} \) can be readily calculated.

First, for a homozygous type \( A_iA_i \), we get
\[ \tilde{C}_{ii}(\xi; M) = \sum_{\xi} C_{ii}(\xi; M) = \frac{\partial \Phi}{\partial \xi_{ii}} (\xi | \xi; M), \]

Consequently, by separating the constant term in the Laurent expansion around the origin of the last quantity regarded as a function of variables \( s_{ab}, t_{ab} (a, b = 1, \ldots, m; a \leq b) \), we obtain for the mean of \( C_{ii} \) an expression
\[ \tilde{C}_{ii}(\xi; M) = \frac{1}{N} \left( F_{ii} + \sum_{\xi_{ii}} \frac{F_{ii}}{2} \right) \left( M_{ii} + \sum_{\xi_{ii}} \frac{M_{ii}}{2} \right). \]

Next, for a heterozygous type \( A_iA_j \), we get
\[ \frac{\partial \Phi}{\partial \xi_{ij}} (\xi | \xi; M) = \frac{1}{N!} \prod_{a,b} \text{Fab} \equiv M_{ab} \equiv N \left( \sum_{a,b} s_{ab} \sum_{a,b} t_{ab} \right)^{N-1} (u_iu_j + u_ju_i), \]
whence follows, similarly as above, for the mean of \( C_{ij} \) an expression
Distributions of Genotypes after a Panmixia

\[ \tilde{C}_{ij} = \sum_{\tilde{C}} C_{ij}(\tilde{C}; \mathcal{M}) = \sum_{C} C_{ij}(C; \mathcal{M}) = \frac{\partial \phi}{\partial z_{ij}} (C; \mathcal{M}) \]

\[ = \frac{1}{N} \left( \left( F_{ii} + \sum_{b \in i} F_{ib} \right) \left( M_{jj} + \sum_{b \in j} M_{jb} \right) \right) \]

\[ + \left( F_{jj} + \sum_{b \in j} F_{jb} \right) \left( M_{ii} + \sum_{b \in i} M_{ib} \right) \]}

Thus, by introducing the quantities defined by

\[ \rho_i^{(F)} = \frac{1}{N} \left( F_{ii} + \sum_{b \in i} F_{ib} \right), \quad \rho_i^{(M)} = \frac{1}{N} \left( M_{ii} + \sum_{b \in i} M_{ib} \right) \quad (i=1, \ldots, m), \]

the expressions for the mean just obtained can be brought into brief forms

\[ \tilde{C}_{ii} = N \rho_i^{(F)} \rho_i^{(M)}, \quad \tilde{C}_{ij} = N (\rho_i^{(F)} \rho_j^{(M)} + \rho_j^{(F)} \rho_i^{(M)}) \quad (i,j=1, \ldots, m; i<j). \]

The last result is quite plausible. In fact, the relative frequencies of the genes \( A_i \) \((i=1, \ldots, m)\) in female-population and in male-population are equal to \( \rho_i^{(F)} \) and \( \rho_i^{(M)} \), respectively. As shown in a previous paper, the random matings between these two populations produce probabilistically a distribution in the next generation with frequencies of \( A_1A_2 \) just given by \( \tilde{C}_{12} \).

4. Variances and covariances.

In order to express the variances, covariances, or, more generally, the quantities concerning the moments of stochastic variables in clearer forms, it will be convenient to introduce an abbreviated notation defined by

\[ [A^n] = \frac{A!}{(A-n)!} = A(A-1) \cdots (A-n+1) \quad (n=0,1, \ldots, A), \]

which represents the number of permutations of \( n \) things taken from \( A \) different things and is often designated by \( A^n \). If we concern

---

merely variances and covariances, it will suffice to introduce the nota-
tion only for \( n=2 \), i.e. \([A_2^2]=A(A-1)\). However, it will become useful
when there concern the moments of higher orders. More generally,
for a given polynomial of any number of arguments
\[
\mathcal{Q}(A_1, \ldots, A_n) = \sum_{\pi} a_{\pi_1} \ldots a_{\pi_n} \prod_{k=1}^{m} A_{\pi_k}^{n_k}
\]
we put
\[
[\mathcal{Q}(A_1, \ldots, A_n)] = \sum_{\pi} a_{\pi_1} \ldots a_{\pi_n} \prod_{k=1}^{m} [A_{\pi_k}^{n_k}]^i.
\]

Now, it is readily shown by induction that there holds for any
derivative of arbitrary order a representation
\[
\prod_{i<j} \left( \frac{\partial}{\partial z_{ij}} \right)^{n_{ij}} \phi(e|s; t) = \frac{1}{N^{2}} \prod_{a<b} \frac{F_{ab}! M_{ab}!}{S_{ab}! P_{ab}} \left[ N \prod_{i<j}^{2} M_{ij}^{-1} \right]^{-1} \sum_{i<j} \sum_{a<b} I_{ab} \cdot \prod_{i} (u_i v_i)^{n_{ij}} \prod_{i<j} (u_i v_j + u_j v_i)^{n_{ij}},
\]
whence follows
\[
\prod_{i<j} \left( \frac{\partial}{\partial z_{ij}} \right)^{n_{ij}} \phi(e) = \frac{1}{[N \prod_{i<j}^{2} M_{ij}]^{1}} \left[ \prod_{i<j} (N \tilde{C})^{n_{ij}} \right]^{i}.
\]

Here the relevant variables in the bracket-notation \([ ]\) for the second
factor are \( m(m+1) \) variables \( F_{ij} \) and \( M_{ij} \) \((i,j=1, \ldots, m; i \leq j)\); as shown
above, we have
\[
N \tilde{C}_{ii} = \left( F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left( M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) \quad (i=1, \ldots, m),
\]
\[
N \tilde{C}_{ij} = \left( F_{ii} + \sum_{b \neq i} \frac{F_{ib}}{2} \right) \left( M_{ij} + \sum_{b \neq j} \frac{M_{ib}}{2} \right) + \left( F_{jj} + \sum_{b \neq j} \frac{F_{jb}}{2} \right) \left( M_{ii} + \sum_{b \neq i} \frac{M_{ib}}{2} \right) \quad (i,j=1, \ldots, m; i < j).
\]

By means of the last relation, any moments of arbitrary order can
be readily computed. As an illustrative example, we shall here deal
with the variances and covariances.

First, the variance of $C_{ab}$ is given by

$$\text{var}(C_{ab}) = \text{var}(C_{ab} | \mathcal{F}, \mathcal{W}) = \sum_{c} (C_{ab} - \tilde{C}_{ab})^2 \psi(c)$$

$$= \sum_{c} \left[ \left( C_{ab} \right) + C_{ab} \right] \psi(c) - \tilde{C}_{ab}$$

$$= \frac{\partial^2 \Phi}{\partial x_{ab}^2} (c) + \tilde{C}_{ab} - \tilde{C}_{ab}$$

$$= (a, b = 1, \ldots, m; a \leq b).$$

Since the value of $\tilde{C}_{ab}$ has been already determined, it is only necessary to substitute the value of a pure derivative of the second order which follows readily from a general formula established above.

Next, the covariance between $C_{ab}$ and $C_{cd}$ is given by

$$\text{cov}(C_{ab}, C_{cd}) = \text{cov}(C_{ab}, C_{cd} | \mathcal{F}, \mathcal{W}) = \sum_{c} (C_{ab} - \tilde{C}_{ab})(C_{cd} - \tilde{C}_{cd}) \psi(c)$$

$$= \sum_{c} C_{ab}C_{cd} \psi(c) - \tilde{C}_{ab}\tilde{C}_{cd} = \frac{\partial^2 \Phi}{\partial x_{ab} \partial x_{cd}} (c) - \tilde{C}_{ab}\tilde{C}_{cd}$$

$$= (a, b, c, d = 1, \ldots, m; a \leq b; c \leq d; A_{ab} \neq A_{cd}).$$

Since the values of $\tilde{C}_{ab}$ and $\tilde{C}_{cd}$ have been already determined, it is only necessary to substitute the value of a mixed derivative of the second order which also follows readily from a general formula established above.

Now, for the sake of completeness, we shall set out below the values of variance as well as covariance of the stochastic variables, after classified according to homozygous and heterozygous types, in more concrete forms.

In general, if the relevant variables in the bracket-notation are $X_{ab} (a, b = 1, \ldots, m, a \leq b; X_{ab} = X_{ba})$, then we get

$$\left[ \left( X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right) \right]^2 = \left( X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right)^2 - \left( X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{4} \right)$$

$$= \left( X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right)^2 \left( X_{jj} + \sum_{b \neq j} \frac{X_{jb}}{2} \right)$$

$$= \left( X_{ii} + \sum_{b \neq i} \frac{X_{ib}}{2} \right) \left( X_{jj} + \sum_{b \neq j} \frac{X_{jb}}{2} \right) - \frac{X_{ij}}{4}$$

$$(i \neq j).$$
Hence, introducing, as before, notations defined by

\[ p_i^{(F)} = \frac{1}{N} \left( F_{ii} + \sum_{\neq i} \frac{F_{ib}}{2} \right) \text{ and } p_i^{(M)} = \frac{1}{N} \left( M_{ii} + \sum_{\neq i} \frac{M_{ib}}{2} \right), \]

we obtain for the values of variance the following expressions:

\[
\text{var}(C_{ii}) = \frac{1}{N(N-1)} \left( N^2 p_i^{(F)^2} - \frac{N p_i^{(F)} + F_{ii}}{2} \right) \left( N^2 p_i^{(M)^2} - \frac{N p_i^{(M)} + M_{ii}}{2} \right) + N p_i^{(F)^2} p_i^{(M)^2} - N^2(p_i^{(F)^2} p_i^{(M)^2})^2
\]

\[ = \frac{N^2}{N-1} \left( p_i^{(F)} p_i^{(M)} \left( 1 - \frac{p_i^{(F)} + p_i^{(M)}}{2} - \frac{1}{N} \right) - \frac{p_i^{(M)^2}}{2} \left( \frac{F_{ii} - p_i^{(F)^2}}{N} \right) \right) - \frac{p_i^{(F)^2}}{2} \left( \frac{M_{ii} - p_i^{(M)^2}}{N} \right) + \frac{1}{4N} \left( F_{ii} - p_i^{(F)} \right) \left( M_{ii} - p_i^{(M)} \right) \right) \}

\[
\text{var}(C_{ij}) = \frac{1}{N(N-1)} \left( \left( N^2 p_j^{(F)^2} - \frac{N p_j^{(F)} + F_{jj}}{2} \right) \left( N^2 p_j^{(M)^2} - \frac{N p_j^{(M)} + M_{jj}}{2} \right) \right)
\]

\[ = \frac{N^2}{N-1} \left( p_j^{(F)} p_j^{(M)} \left( 1 - \frac{p_j^{(F)} + p_j^{(M)}}{2} - \frac{1}{N} \right) - \frac{p_j^{(M)^2}}{2} \left( \frac{F_{jj} - p_j^{(F)^2}}{N} \right) \right)
\]

\[ + \frac{1}{4N} \left( F_{ij} - p_j^{(F)} \right) \left( M_{ij} - p_j^{(M)} \right) \left( F_{jj} - p_j^{(F)^2} \right) \left( M_{jj} - p_j^{(M)^2} \right) \right) + \frac{N(p_j^{(F)^2} p_j^{(M)^2} + p_j^{(F)^2} p_j^{(M)^2})^2}
\]

\[ = \frac{N^2}{N-1} \left( p_j^{(F)} p_j^{(M)} \left( 1 - \frac{p_j^{(F)} + p_j^{(M)}}{2} - \frac{1}{N} \right) - \frac{p_j^{(M)^2}}{2} \left( \frac{F_{jj} - p_j^{(F)^2}}{N} \right) \right)
\]

\[ + \frac{1}{4N} \left( F_{ij} - p_j^{(F)} \right) \left( M_{ij} - p_j^{(M)} \right) \left( F_{jj} - p_j^{(F)^2} \right) \left( M_{jj} - p_j^{(M)^2} \right) \right) + \frac{N(p_j^{(F)^2} p_j^{(M)^2} + p_j^{(M)^2} p_j^{(F)^2})^2}
\]
Distributions of Genotypes after a Panmixia

\[ + \frac{1}{4N} \left( \left( \frac{F_{ii}}{N} + p_{i}^{(P)} \right) \left( \frac{M_{jj}}{N} + p_{j}^{(M)} \right) + \frac{1}{2} \frac{F_{jj}}{N} \frac{M_{jj}}{N} \right) \]

\[ + \left( \frac{F_{ij}}{N} + p_{i}^{(M)} \right) \left( \frac{M_{ij}}{N} + p_{j}^{(P)} \right) \] \quad \text{for } i \neq j.

Similarly, we obtain for the values of covariance the following expressions:

\[
\text{cov} (C_{ii}, C_{ik}) = \frac{1}{N(N-1)} \left( \left( N^{2}p_{i}^{(P)} - \frac{F_{ik}}{4} \right) \left( N^{2}p_{k}^{(M)}p_{i}^{(M)} - \frac{M_{ik}}{4} \right) 
+ \left( N^{2}p_{i}^{(P)}p_{k}^{(M)} \right) - \frac{F_{ik}}{4} \left( N^{2}p_{i}^{(M)}p_{k}^{(M)} - \frac{M_{ik}}{4} \right) \right)
\]

\[
+ \frac{1}{8N} \left( \left( \frac{F_{ii}}{N} + p_{i}^{(P)} \right) \frac{M_{ik}}{N} + \left( \frac{M_{ii}}{N} + p_{i}^{(M)} \right) \frac{F_{ik}}{N} \right) \] \quad \text{for } k \neq i.

\[
\text{cov} (C_{ii}, C_{ih}) = \frac{1}{N(N-1)} \left( \left( N^{2}p_{i}^{(P)}p_{h}^{(F)} - \frac{F_{ih}}{4} \right) \left( N^{2}p_{h}^{(M)}p_{i}^{(M)} - \frac{M_{ih}}{4} \right) 
- N^{2}p_{i}^{(P)}p_{h}^{(M)}p_{i}^{(F)}p_{h}^{(M)} \right)
\]

\[
+ \frac{1}{16N} \left( \frac{F_{ih}}{N} \frac{M_{ih}}{N} \right) \] \quad \text{for } k \neq i.

\[
\text{cov} (C_{ii}, C_{hk}) = \frac{1}{N(N-1)} \left( \left( N^{2}p_{i}^{(P)}p_{h}^{(F)} - \frac{F_{ih}}{4} \right) \left( N^{2}p_{h}^{(M)}p_{i}^{(M)} - \frac{M_{ih}}{4} \right) 
- N^{2}p_{i}^{(P)}p_{h}^{(M)}p_{i}^{(F)}p_{h}^{(M)} \right)
\]

\[
+ \left( N^{2}p_{i}^{(P)}p_{h}^{(F)} \right) \left( N^{2}p_{i}^{(M)}p_{h}^{(M)} - \frac{M_{ih}}{4} \right) \] \quad \text{for } k \neq i.
\[
- \frac{p_i^{(F)}}{4} \left( p_k^{(F)} \frac{M_{ih}}{N} + p_{ik}^{(F)} \frac{M_{ih}}{N} \right) + \frac{1}{16N} \left( \frac{F_{ih}}{N} \frac{M_{ih}}{N} + \frac{F_{ik}}{N} \frac{M_{ih}}{N} \right) \]

\[(h, k = i; h \neq k),\]

\[
\text{cov}(C_{ij}, C_{ik}) = \frac{1}{N(N-1)} \left\{ \left( N^2 p_i^{(F)} p_j^{(F)} + \frac{Np_i^{(F)} + F_{ii}}{2} \right) \left( N^2 p_j^{(M)} p_k^{(M)} + \frac{M_{jk}}{4} \right) \right. \]

\[+ \left( N^2 p_i^{(F)} p_j^{(F)} - \frac{F_{ij}}{4} \right) \left( N^2 p_j^{(M)} p_k^{(M)} + \frac{M_{ij}}{4} \right) \]

\[+ \left( N^2 p_i^{(F)} p_j^{(F)} - \frac{F_{j}}{4} \right) \left( N^2 p_i^{(M)} p_k^{(M)} + \frac{M_{ik}}{4} \right) \]

\[= \frac{N^2}{N-1} \left\{ \left( \frac{p_i^{(F)} p_j^{(M)}}{4} + \frac{p_j^{(F)} p_i^{(M)}}{4} \right) \left( \frac{p_i^{(F)} p_j^{(M)}}{4} + \frac{p_j^{(F)} p_i^{(M)}}{4} \right) \right. \]

\[+ \frac{1}{N} \left( \frac{F_{ij}}{N} + \frac{F_{ji}}{N} \right) \left( \frac{M_{ij}}{N} + \frac{M_{ji}}{N} \right) \left( i \neq j; k = i, j \right), \]

\[
\text{cov}(C_{ij}, C_{kk}) = \frac{1}{N(N-1)} \left\{ \left( N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left( N^2 p_j^{(M)} p_k^{(M)} - \frac{M_{jk}}{4} \right) \right. \]

\[+ \left( N^2 p_i^{(F)} p_k^{(F)} - \frac{F_{ik}}{4} \right) \left( N^2 p_j^{(M)} p_k^{(M)} - \frac{M_{jk}}{4} \right) \]

\[+ \left( N^2 p_j^{(F)} p_k^{(F)} - \frac{F_{jk}}{4} \right) \left( N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \]

\[+ \left( N^2 p_j^{(F)} p_k^{(F)} - \frac{F_{jk}}{4} \right) \left( N^2 p_i^{(M)} p_k^{(M)} - \frac{M_{ik}}{4} \right) \]
Distributions of Genotypes after a Panmixia

\[-N^2(\sum_i p_i^{(F)}p_i^{(M)} + \sum_j p_j^{(F)}p_j^{(M)})(\sum_k p_k^{(F)}p_k^{(M)} + \sum_l p_l^{(F)}p_l^{(M)})\]

\[= \frac{N}{N-1} \left\{ \left( \sum_i p_i^{(F)}p_i^{(M)} + \sum_j p_j^{(F)}p_j^{(M)} \right) \left( \sum_k p_k^{(F)}p_k^{(M)} + \sum_l p_l^{(F)}p_l^{(M)} \right) \right\} \]

\[+ \sum_i p_i^{(F)} \left( \frac{M_{ih}}{N} + \frac{F_{ih}}{N} \right) + \sum_j p_j^{(F)} \left( \frac{M_{jh}}{N} + \frac{F_{jh}}{N} \right) \]

\[+ \sum_k p_k^{(F)} \left( \frac{M_{ih}}{N} + \frac{F_{ih}}{N} \right) + \sum_l p_l^{(F)} \left( \frac{M_{jh}}{N} + \frac{F_{jh}}{N} \right) \]

\[+ \frac{1}{16N} \left( \frac{F_{ih}}{N} \frac{M_{jh}}{N} + \frac{F_{ih}}{N} \frac{M_{jh}}{N} + \frac{F_{jh}}{N} \frac{M_{ih}}{N} + \frac{F_{jh}}{N} \frac{M_{ih}}{N} \right) \]

\[(i \neq j; h \neq k; h, k \neq i, j). \]

It would be noted, in passing, that there holds an identical relation of dependence, i.e.

\[0 = \sum_{ab} \left( \sum_{cd} (C_{ab} - \bar{C}_{ab}) \right)^2 \var(\xi) = \sum_{ab} \var(C_{ab}) + 2 \sum_{a \neq b} \cov(C_{ab}, C_{cd}); \]

the last summation extends over all the possible pairs \((ab, cd)\).

Asymptotic behaviors of variance and covariance as \(N\) increases can be readily deduced from the expressions just established. In fact, since there hold always \(F_{ab}/N \leq 1\) and \(M_{ab}/N \leq 1\), we obtain

\[\var(C_{ii}) = N \left\{ p_i^{(F)}p_i^{(M)} \left( 1 - \frac{p_i^{(F)} + p_i^{(M)}}{2} \right) \right. \]

\[- \frac{p_i^{(M,2)}}{2N} \left( \frac{M_{ii}}{N} - p_i^{(M,2)} \right) \left. - \frac{p_i^{(F,2)}}{2N} \left( M_{ii} - p_i^{(F,2)} \right) \right\} + O(1), \text{ etc.} \]

On the other hand, if the original distributions \(\mathcal{F}\) and \(\mathcal{M}\) show, in particular, the same equilibrium state, i.e., when there hold

\[F_{bb} = M_{bb} = Np_b^2, \quad F_{ab} = M_{ab} = N2p_ap_b \quad (a, b = 1, \ldots, m; a < b), \]

the expressions may be reduced to fairly simple forms. Namely, we then get

\[\var(C_{ii}) = \frac{N^2}{N-1} p_i^2 \left( 1 - p_i \left( 1 - \frac{3 + p_i}{3N} \right) \right), \]
\[
\text{var}(C_{ij}) = \frac{N^2}{N-1} p_i p_j \left( 2 - p_i - p_j - \frac{3 - p_i - p_j - 2p_i p_j}{2N} \right) \quad (i \neq j);
\]

\[
\text{cov}(C_{ii}, C_{hk}) = -\frac{N^3}{N-1} p_i^2 p_h \left( 1 - \frac{1 + p_i}{2N} \right) \quad (k \neq i),
\]

\[
\text{cov}(C_{ii}, C_{hh}) = \frac{1}{4} \frac{N}{N-1} p_i^2 p_h^2 \quad (k \neq i),
\]

\[
\text{cov}(C_{ij}, C_{ij}) = -\frac{N^3}{N-1} p_i p_j p_h \left( 1 - 2p_i - \frac{1 + 2p_i}{2N} \right) \quad (i \neq j; k \neq i, j),
\]

\[
\text{cov}(C_{ij}, C_{hh}) = \frac{N}{N-1} p_i p_j p_h p_h \quad (i \neq j; h, k \neq i, j; h \neq k).
\]

Thus, in this particular case, \( \text{cov}(C_{ii}, C_{hh}) \), \( \text{cov}(C_{ii}, C_{hh}) \) and \( \text{cov}(C_{ij}, C_{hh}) \) are really of order \( O(1) \), i.e. bounded regardless of the values of \( N > 1 \), while the remaining \( \text{cov} \)'s as well as the \( \text{var} \)'s are, in general, unbounded with order \( O(N) \).

Department of Mathematics,
Tokyo Institute of Technology.